AWARD NUMBER: W81XWH-16-1-0797

TITLE: An Adaptive Tutor for Improving Visual Diagnosis

PRINCIPAL INVESTIGATOR: Martin V. Pusic, MD

RECIPIENT: New York University
New York, NY 10016

REPORT DATE: October 2017

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2017	Annual	30 Sep 2016 - 29 Sep 2017

4. TITLE AND SUBTITLE
An Adaptive Tutor for Improving Visual Diagnosis

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0797

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Martin V. Pusic – NYUSoM; David Cook – Mayo Clinic; Rose Hatala – UBC; Matt Lineberry – KUMC.

E-Mail: Martin.Pusic@nyumc.org; Cook.David33@mayo.edu; rhatala@mac.com; mlineberry@kumc.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NYU School of Medicine

550 First Ave.,
New York, NY 10016

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT Online cognitive trainers for visual diagnosis can transcend institutional barriers to enable broad distribution of learning material. However, most current examples are based on declarative knowledge instructional designs that deliver outcomes that are only indirectly connected to patient care. Our key contention is that cognitive learning platforms, using evidence-based instructional designs, can facilitate efficient and effective visual diagnosis skill development and maintenance. **Progress to date:** We have had success in three key components of the eventual adaptive tutor: 1) we have assembled a corpus of 80,000 ECGs with their associated clinical information and have organized that corpus into 2) a prototype presentation database that allows any stakeholder to search and download ECGs according to any of the 120 American Heart Association diagnostic labels. 3) We have completed three pilot studies designed to inform the design of the adaptive tutor including a) focus groups to develop a relative "importance" ranking, b) pairwise comparisons by cardiologists to determine the feasibility of complexity ranking of ECGs and c) exploring the degree to which two overlapping ECG phenotypes can be confused and how this can be statistically modeled. We are now well positioned to use these materials and methods to carry out the next phase, a large prospective data collection and subsequent impact trial of adaptive learning.

15. SUBJECT TERMS
Electrocardiogram (ECG), ECG standards, medical education, visual pattern recognition, educational models, computer-assisted instructions, clinical competence, continuing medical education, psychometrics, e-learning, simulation, and instructional design.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	26	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. **INTRODUCTION:** Narrative that briefly (one paragraph) describes the subject, purpose and scope of the research.

Online cognitive trainers for visual diagnosis can transcend institutional barriers to enable broad distribution of learning material. However, most current examples are based on declarative knowledge instructional designs that deliver outcomes that are only indirectly connected to patient care. Our key contention is that cognitive learning platforms, using evidence-based instructional designs, can facilitate efficient and effective visual diagnosis skill development and maintenance. Our objective is to create an adaptive tutor that allows any recruit to immediately begin classifying authentic cases. The tutor would determine their baseline proficiency within a minimum of cases, and then systematically tailors case-based learning until s/he achieves a reliable pre-transition competency, the latter based on cases directly relevant to the context to which the recruit will be deployed. **Specific Aim: 1.** To assemble a massive online ECG library from authentic field cases collected from a representative clinical setting: the Emergency Department; **SA2.** To develop both ontologic and statistical models of the ECG cases so as to inform the rational design of the adaptive learning system; **SA3.** To develop an evidence-based "learning adaptation algorithm" that can ensure efficient and reliable development of skill at scale.

2. **KEYWORDS:** Provide a brief list of keywords (limit to 20 words).

Electrocardiogram (ECG), ECG standards, diagnosis, medical education, visual pattern recognition, educational models, computer-assisted instruction/methods, reaction time, technology, diagnostic errors, clinical competence, continuing medical education, psychometrics/statistics and numerical data, computer-assisted diagnosis, interpretation difficulty, e-learning, simulation, and instructional design.

3. **ACCOMPLISHMENTS:** The PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency Grants Officer whenever there are significant changes in the project or its direction.

What were the major goals of the project?
List the major goals of the project as stated in the approved SOW. If the application listed milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

Specific Aims Yr. 1 timeline includes:
SA 1: To assemble a massive online ECG library from authentic field cases collected from a representative clinical setting: the Emergency Department
SA 1: Task 1) USAMRMC ORP approval Obtained – (Completed - December 19, 2016) – Pusic, NYU; Cook, Mayo Clinic; Hatala, UBC; Lineberry, KUMC.
SA 1: Task 2) Preliminary labeling of full set of ECGs – (95%) – Pusic, NYU; and Lineberry, KUMC.
SA 1: Task 3) Download ECG cases from MUSE and EPIC systems – (Completed) – Pusic, NYU.
SA 1: Task 4) Focus Group for assigning cognitive task scores, difficulty ranking of diagnostic labels – (Completed) – Pusic, NYU.
SA 1: Task 5) Programing of presentation software – (90%) – Pusic, NYU
SA 1: Task 6) IRB approvals at all 4 site (NYU, KUMC, UBC, and Mayo Clinic) – (Completed) - Pusic, NYU; Cook, Mayo Clinic; Hatala, UBC; Lineberry, KUMC.

SA 2: To develop both ontologic and statistical models of the ECG cases so as to inform the rational design of the adaptive learning system
SA 2: **Task 1**) Natural language processing across the ECG set to identify full diagnostic labeling – (90%) – Pusic, NYU; and Lineberry, KUMC.
SA 2: Task 2) Pilot Data Collection – recruit 80 clinicians – (On Hold*: 60; Completed data collection for Focus Group #1 (16 actual/10 planned); and Confusable diagnosis study 36/20) – Pusic, NYU; Cook, Mayo Clinic; Hatala, UBC.
SA 2: Task 3) Pairwise Comparisons – recruited 20 cardiology fellows/residents – (Completed N=20/20) – Pusic, NYU.
SA 2: Task 4) Case Characteristics Model Initial Calibration – (Ongoing) – Pusic, NYU; and Lineberry, KUMC.
SA 2: Task 5) Programming Study Database – (Ongoing) - Pusic, NYU; and Lineberry, KUMC.
SA2: Task 6) Main Cohort Study - (Start) - Pusic, NYU; Cook, Mayo Clinic; Hatala, UBC.
SA 3: To develop an evidence-based "learning adaptation algorithm" that can ensure efficient and reliable development of skill at scale

For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of stated goals not met. Description shall include pertinent data and graphs in sufficient detail to explain any significant results achieved. A succinct description of the methodology used shall be provided. As the project progresses to completion, the emphasis in reporting in this section should shift from reporting activities to reporting accomplishments.

1) Major activities; and 2) specific objectives
Specific Aims timeline for Yr. 1 includes:
SA 1: To assemble a massive online ECG library from authentic field cases collected from a representative clinical setting: the Emergency Department – (Ongoing aim – 95%)
SA 1: Task 1) Submit documents for USAMRMC ORP approval and IRB applications – (Completed. USAMRMC ORP approved NYU site with all the necessary documentation. Approvals for studies took place in Q2 and Q3. Applications for IRB were submitted for all the participating sites for studies that take place in Q4 and beyond. NYU IRB was fully approved on October 14, 2016; University of British Columbia IRB was fully approved on March 9, 2017; KUMC site fully approved April 7th; and Mayo Clinic IRB was fully approval on May 1, 2017. HRPO completed approval after revisions to Mayo protocol. No delays to study recruitment are anticipated. Pusic, NYU; Cook, Mayo Clinic; Hatala, UBC; and Lineberry, KUMC.
SA 1: Milestone 1) USAMRMC ORP approval Obtained – (After review of the necessary documents USAMRMC ORP approval was obtained on December 19, 2016 for Part I of the protocol: Part II approved on May 16, 2017) – Pusic, NYU.
SA 1: Task 2) Download ECG cases from MUSE and EPIC Electronic Health Records, initial labeling using reporting functions – After solving technical issues 95,000 candidate ECG cases were downloaded; the EPIC collateral data has also been downloaded. We are working to validate the merge of the two datasets resulting in a bank of 80,000 (augmented from the planned 20,000). Pusic, NYU.
SA 1: Task 3) Focus groups for assigning cognitive task scores, difficulty ranking of 120 American Heart Association (AHA) diagnostic labels. We divided this into three sessions, one with cardiologists, second with emergency medicine and internal medicine clinicians, and group of nurse practitioners. First focus group with cardiologist was held on March 23, 2016. While, second was held April 6, 2016. Third group of nurses were spread out to be one on one session due to their availability of time and location. Transcription was completed end of May for first two groups, while nurse interviews transcriptions are in progress. Importance ranking has been captured and is being entered for each case in the bank for all groups. Coding of transcription is in progress. Pusic, NYU.

SA 2: To develop both ontologic and statistical models of the ECG cases so as to inform the rational design of the adaptive learning system – (Ongoing aim)
SA 2: Task 1) Natural language processing across the ECG set to identify full diagnostic labeling – all 80,000 cases have been parsed by AHA diagnostic category. We are in the process of validating that parsing both with the original ECG and now with the downloaded EMR data. Final outcome will be a fully categorized set. Pusic, NYU; and Lineberry, KUMC.
SA 2: Task 2) Pilot Data Collection – recruit 80 clinicians – recruitment in process

- Confusable diagnoses –recruited xx experts and xx novices (total 36) who completed the study maneuver. Preliminary results show the anticipated pattern of experts being able to distinguish 5 thresholds of certainty for the Pericarditis/STEMI distinction. Pusic, NYU
- Reaction Time – plan is to recruit 60 subjects; we have put this on hold until we analyze the pairwise comparison data fully to better frame the reaction time metric. We would like to get a reaction time metric that takes into account the cognitive complexity of each case. We would do that best if we do this based on the pairwise comparison data.

SA 2: Task 3) Pairwise Comparisons – recruit 20 cardiology fellows/residents- (20/20 completed)

- Started initial analysis of the pairwise data. Pusic, NYU

SA 2: Task 4) Case Characteristics Model Initial Calibration – (Ongoing) - Pusic, NYU; and Lineberry, KUMC. Target study populations all identified.
SA 2: Task 5) Programming Study Database – (Ongoing) - Pusic, NYU; and Lineberry, KUMC.
SA2: Task 6) Main Cohort Study - (Ongoing) - Pusic, NYU; Cook, Mayo Clinic; Hatala, UBC.

3) Significant results or key outcomes, including major findings, developments, or conclusions

D.4.3.4. Difficulty by Pairwise Comparisons. (Appendix 1).

Twenty cardiology fellows (NYU Fellows) each ranked 12 ECGs according to their perceived difficulty in a pairwise comparison process that had them compare every ECG with every other one, choosing one as being more difficult. They also listed their reason for choosing one over another. This was designed to inform the cognitive complexity ranking shown in D.4.2.1. Our findings were surprising in that the complexity ranking correlated only minimally with our *a priori* listed cognitive complexities based on the documented visual features of the ECGs. A building theme in this research is that it is difficult for an expert to examine an ECG and predict its difficulty for others.

D.4.3.5. Rankings of Two Confusable Diagnoses. (Appendix 2).

We had 36 subjects (18 experts; 18 novices) rate 20 ECGs (10 Pericarditis; 10 ST elevation MI) in order for us to model the "confusability" of these two important diagnoses. We were able to fit a five-choice polytomous IRT model using a Markov-Chain Monte-Carlo Bayesian technique. This model was able to show the ability, and limitations, of expert and novice EM physicians and cardiologists to classify ECGs as to the likelihood of ST Elevation Myocardial Infarction.

D.4.3.4. Importance Rankings of ECG Diagnoses by Expert Group. (Appendix 3).

Using two focus groups and several semi-structured interviews, we collected the views of experts in three disciplines (Emergency Medicine, Cardiology and Nursing) as to which diagnostic categories are most important for a practicing clinicians. Each participant ordered the 120 American Heart Association diagnostic labels as to their importance. The results will guide item selection for the educational intervention.

What opportunities for training and professional development has the project provided?
If the project was not intended to provide training and professional development opportunities or there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe opportunities for training and professional development provided to anyone who worked on the project or anyone who was involved in the activities supported by the project. "Training" activities are those in which individuals with advanced professional skills and experience assist others in attaining greater proficiency. Training activities may include, for example, courses or one-on-one work with a mentor. "Professional development" activities result in increased knowledge or skill in one's area of expertise and may include workshops, conferences, seminars, study groups, and individual study. Include participation in conferences, workshops, and seminars not listed under major activities.

Study PI (Martin Pusic - NYU), co-investigator (Matthew Lineberry - KUMC) and NYU staff members (Jacqueline Gutman, So-Young Oh, Eric Feng, and Sidrah Malik) attended a day long seminar on April 21, 2017 at NYU School of Medicine, where the focus was to explore the use of adaptive learning techniques in visual diagnosis with a view to optimizing both individual and group- level learning. We had the opportunity to learn from Dr. Chris Gibbons, Director of Health Assessment and Innovation at the Psychometrics Center of University of Cambridge along with Drs. Alyssa Wise, Yoav Bergner, and Peter Halpin, experts in learning sciences and Educational technology, and Applied Statistics at NYU. These techniques, and in particular Computer Adaptive Design, will help us develop our future sub-studies and applications.

How were the results disseminated to communities of interest?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the results were disseminated to communities of interest. Include any outreach activities that were undertaken to reach members of communities who are not usually aware of these project activities, for the purpose of enhancing public understanding and increasing interest in learning and careers in science, technology, and the humanities.

The ECG project database was used by the NYU Entrepreneurial Institute as the basis of one of the Healthcare Challenges during the 2017 NYU Healthcare Makerthon. Interdisciplinary teams of students from healthcare, computer science and the liberal arts combined forces to reimagine the visualization of the ECG using modern tools. The competition is ongoing. Details are available at: http://entrepreneur.nyu.edu/resource/healthcare-makerthon/

Describe briefly what you plan to do during the next reporting period to accomplish the goals and objectives.

Specific Aims timeline for Yr. 2 Q 1 includes:
SA 2: Task 6) Main cohort study – (Ongoing) - Pusic, NYU; Cook, Mayo Clinic; Hatala, UBC.

- We are in the final stages of developing the specific presentation algorithm for the prospective cohort study of 500 learners completing at least 100 ECGs each. The design of the prospective cohort study will be based on information from each of the sub-studies completed to date.

SA 3: To develop an evidence-based "learning adaptation algorithm" that can ensure efficient and reliable development of skill at scale. This will in turn build on the results of the data collected in the prospective cohort study.

4. **IMPACT:** Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:

What was the impact on the development of the principal discipline(s) of the project?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research in the principal disciplinary field(s) of the project. Summarize using language that an intelligent lay audience can understand (Scientific American style).

Nothing to Report.

What was the impact on other disciplines?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

Nothing to Report.

What was the impact on technology transfer?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use, including:

- *transfer of results to entities in government or industry;*
- *instances where the research has led to the initiation of a start-up company; or*
- *adoption of new practices.*

Nothing to Report.

What was the impact on society beyond science and technology?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

- *improving public knowledge, attitudes, skills, and abilities;*
- *changing behavior, practices, decision making, policies (including regulatory policies), or social actions; or*
- *improving social, economic, civic, or environmental conditions.*

Nothing to Report.

5. **CHANGES/PROBLEMS:** The Project Director/Principal Investigator (PD/PI) is reminded that the recipient organization is required to obtain prior written approval from the awarding agency Grants Officer whenever there are significant changes in the project or its direction. If not previously reported in writing, provide the following additional information or state, "Nothing to Report," if applicable:

Changes in approach and reasons for change
Describe any changes in approach during the reporting period and reasons for these changes. Remember that significant changes in objectives and scope require prior approval of the agency.

Sub-study D.4.3.6. "Reading Time as a Proxy Neuroplasticity Measure"
The study team felt that carrying out this study prior to having the results of the Pairwise Comparison study would be potentially problematic as the choice of cases for determining "reaction time" was to be based on the cognitive complexity ranking. Once we have fully analyzed the pairwise comparison data (D.4.3.4.) and the prospective cohort data (D.6.1.), we will be well positioned to carry out the full "Reading time" study as originally intended. Thus, we will accomplish all studies within the overall timeframe, having altered only the sequence in which this "Reading Time" study is carried out relative to the others, preferring to benefit from the available reading time data.

Actual or anticipated problems or delays and actions or plans to resolve them
Describe problems or delays encountered during the reporting period and actions or plans to resolve them.

Initial delay of processing IRB at collaborating sites was eventually resolved such that all sites were fully approved by IRBs by May 1, 2017. This did not delay recruitment since recruitment is not required for earlier sub studies. Delay in hiring delayed our start or other sub-studies and programming of the ECG software. We have also worked through some technical issues related to obtaining full and complete "atomic" XML versions of the ECGs but those have been surmounted such that we now have a full set of ECGs as planned. For next quarter, we anticipate largely completing the assembly of the ECG bank with both ECG-data and collateral clinical data from our EPIC electronic health record. We will report the specificity with which we are able to algorithmically label all 80,000 ECGs through the use of random subsamples reviewed by the investigators. We also anticipate completing the NYU pilot studies (enrollment 100). We do not anticipate any problems with recruiting the proposed number of participants or having received completed data from them. We plan to complete the analysis for NYU pilot studies and reporting of initial results.

Changes that had a significant impact on expenditures
Describe changes during the reporting period that may have had a significant impact on expenditures, for example, delays in hiring staff or favorable developments that enable meeting objectives at less cost than anticipated.

Challenge of delay in hiring staff for the grant. Several candidates were interviewed for the few positions list on our grant. We have successfully hired the project manager (Ms. Malik) of the grant in December. We carried out a 2-month process to hire a programmer (Mr. Feng) which was successfully completed January 8th, 2017. Instructional designer (Ms. Oh) for grant started working from quarter 2 and statistician (Ms. Gutman) start by quarter 3. Additionally, we have a balance of 35.26% of total costs much of this is through unliquidated obligations for subcontracts; currently we have a total of $118,563.65. Pending invoices from UBC ($23,706), KUMC ($65,637), and Mayo Clinic ($29,220.65). We expect to receive full and final invoicing for UBC and Mayo Clinic within 60 days of budget end date which will be November 29, 2017. KUMC has requested additional time to fully expend and invoice for Dr. Lineberry's effort. If all unliquidated obligations are deducted from the balance we would expect carryover for only 3.83% with a balance of $17,891.93.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents
Describe significant deviations, unexpected outcomes, or changes in approved protocols for the use or care of human subjects, vertebrate animals, biohazards, and/or select agents during the reporting period. If required, were these changes approved by the applicable institution committee (or equivalent) and reported to the agency? Also specify the applicable Institutional Review Board/Institutional Animal Care and Use Committee approval dates.

Significant changes in use or care of human subjects

Nothing to Report.

Significant changes in use or care of vertebrate animals.

Nothing to Report.

Significant changes in use of biohazards and/or select agents

Nothing to Report.

6. **PRODUCTS:** List any products resulting from the project during the reporting period. If there is nothing to report under a particular item, state "Nothing to Report."

- **Publications, conference papers, and presentations**
  Report only the major publication(s) resulting from the work under this award.

  Journal publications. *List peer-reviewed articles or papers appearing in scientific, technical, or professional journals. Identify for each publication: Author(s); title; journal; volume: year; page numbers; status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

  Nothing to Report.

  Books or other non-periodical, one-time publications. *Report any book, monograph, dissertation, abstract, or the like published as or in a separate publication, rather than a periodical or series. Include any significant publication in the proceedings of a one-time conference or in the report of a one-time study, commission, or the like. Identify for each one-time publication: Author(s); title; editor; title of collection, if applicable; bibliographic information; year; type of publication (e.g., book, thesis or dissertation); status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

  Nothing to Report.

  Other publications, conference papers, and presentations. *Identify any other publications, conference papers and/or presentations not reported above. Specify the status of the publication as noted above. List presentations made during the last year (international, national, local societies, military meetings, etc.). Use an asterisk (*) if presentation produced a manuscript.*

  1) International Outcome Measurement Conference (IOMC) 2017 – Presentation
     a. Abstract was successfully submitted for presentation on September 15, 2017 in Chicago, IL. Presentation title "A Graded Response Model for Diagnostic Discernment in Visual Diagnosis". We presented early findings of our Confusability study.

- **Website(s) or other Internet site(s)**
  List the URL for any Internet site(s) that disseminates the results of the research activities. A short description of each site should be provided. It is not necessary to include the publications already specified above in this section.

  Nothing to Report.

- **Technologies or techniques**
  Identify technologies or techniques that resulted from the research activities. In addition to a description of the technologies or techniques, describe how they will be shared.

  Nothing to Report.

- **Inventions, patent applications, and/or licenses**
  Identify inventions, patent applications with date, and/or licenses that have resulted from the research. State whether an application is provisional or non-provisional and indicate the application number. Submission of this information as part of an interim research performance progress report is not a substitute for any other invention reporting required under the terms and conditions of an award.

  Nothing to Report.

- **Other Products**
  Identify any other reportable outcomes that were developed under this project. Reportable outcomes are defined as a research result that is or relates to a product, scientific advance, or research tool that makes a meaningful contribution toward the understanding, prevention, diagnosis, prognosis, treatment, and/or rehabilitation of a disease, injury or condition, or to improve the quality of life. Examples include:
  - *data or databases;*
  - *biospecimen collections;*
  - *audio or video products;*
  - *software;*
  - *models;*
  - *educational aids or curricula;*
  - *instruments or equipment;*
  - *research material (e.g., Germplasm; cell lines, DNA probes, animal models);*
  - *clinical interventions;*
  - *new business creation; and*
  - *other.*

  Nothing to Report.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Provide the following information for: (1) PDs/PIs; and (2) each person who has worked at least one person month per year on the project during the reporting period, regardless of the source of compensation (a person month equals approximately 160 hours of effort). If information is unchanged from a previous submission, provide the name only and indicate "no change."

Example:

Name:	*Mary Smith*
Project Role:	*Graduate Student*
Researcher Identifier (e.g. ORCID ID):	*1234567*
Nearest person month worked:	*5*
Contribution to Project:	*Ms. Smith has performed work in the area of combined error-control and constrained coding.*
Funding Support:	*The Ford Foundation (Complete only if the funding support is provided from other than this award).*

Name:	Martin V. Pusic, MD
Project Role:	Principal Investigator
Contribution to Project:	No Change
Name:	David Cook, MD
Project Role:	Co-Investigator
Contribution to Project:	No Change
Name:	Rose Hatala, MD
Project Role:	Co-Investigator
Contribution to Project:	No Change
Name:	Matthew Lineberry, PhD
Project Role:	Co-Investigator
Contribution to Project:	No Change
Name:	Marc Triola, MD
Project Role:	Co-Investigator
Contribution to Project:	No Change
Name:	Silas Smith, MD
Project Role:	Co-Investigator
Contribution to Project:	No Change
Name:	Jennifer Hill, PhD, PRIISM
Project Role:	Co-Investigator
Contribution to Project:	No Change

Name:	Barry Rosenzweig, MD
Project Role:	Co-Investigator
Contribution to Project:	No Change
Name:	Greta Elysee
Project Role:	Program Coordinator
Contribution to Project:	No Change
Name:	Sidrah Malik
Project Role:	Project Manager
Contribution to Project:	No Change
Name:	Eric Feng
Project Role:	Programmer
Contribution to Project:	No Change
Name:	So–Young Oh
Project Role:	Instructional Designer
Contribution to Project:	No Change
Name:	Jacqueline Gutman
Project Role:	Statistician
Contribution to Project:	No Change

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

If the active support has changed for the PD/PI(s) or senior/key personnel, then describe what the change has been. Changes may occur, for example, if a previously active grant has closed and/or if a previously pending grant is now active. Annotate this information so it is clear what has changed from the previous submission. Submission of other support information is not necessary for pending changes or for changes in the level of effort for active support reported previously. The awarding agency may require prior written approval if a change in active other support significantly impacts the effort on the project that is the subject of the project report.

Nothing to Report.

What other organizations were involved as partners?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe partner organizations – academic institutions, other nonprofits, industrial or commercial firms, state or local governments, schools or school systems, or other organizations (foreign or domestic) – that were involved with the project. Partner organizations may have provided financial or in-kind support, supplied facilities or equipment, collaborated in the research, exchanged personnel, or otherwise contributed.
Provide the following information for each partnership:
Organization Name:
Location of Organization: (if foreign location list country)
Partner's contribution to the project (identify one or more)

- *Financial support;*
- *In-kind support (e.g., partner makes software, computers, equipment, etc., available to project staff);*
- *Facilities (e.g., project staff use the partner's facilities for project activities);*
- *Collaboration (e.g., partner's staff work with project staff on the project);*
- *Personnel exchanges (e.g., project staff and/or partner's staff use each other's facilities, work at each other's site); and*
- *Other.*

Organization Name:	NYU School of Medicine
Location of Organization:	New York, NY
Partner's contribution to the Project:	DoD Grant Recipient
Organization Name:	Mayo Clinic
Location of Organization:	Rochester, Minnesota
Partner's contribution to the Project:	Collaboration
Organization Name:	University of British Columbia
Location of Organization:	Vancouver, Canada
Partner's contribution to the Project:	Collaboration
Organization Name:	University of Kansas Medical Center
Location of Organization:	Kansas City, Kansas
Partner's contribution to the Project:	Collaboration

8. SPECIAL REPORTING REQUIREMENTS

QUAD CHART: Please see below.

An Adaptive Tutor for Improving Visual Learning

ERMS/Log Number: DM160044; **Task Title:** Adaptive Tutor Using Methodologies for Neuroplasticity
Award Number: W81XWH-16-1-0797

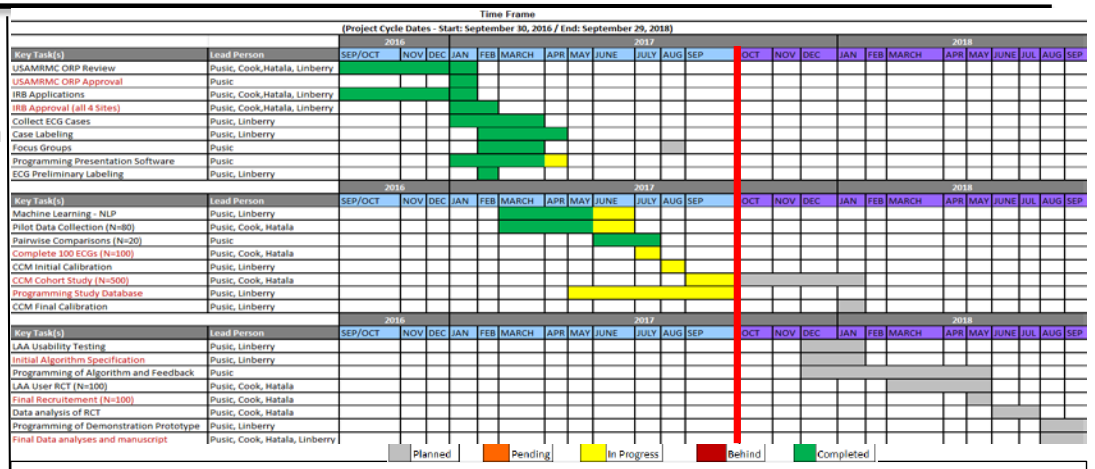
PI: Martin Victor Pusic **Org:** NYU School of Medicine **Award Amount:** $1,477,721

Study/Product Aim(s)

- Assemble and label an authentic library of digital ECGs from Hospital Information System for deliberate practice by clinicians.
- Develop "Case Characteristics Model" (CCM) that can algorithmically assign difficulty and importance indices to cases.
- Develop "Learning Adaptation Algorithm" (LAA) that chooses cases from ECG library based on an individual's learning curve.

Approach

- We would harvest and label 20,000 ECG cases from hospital EHR including images, clinical and laboratory data and outcome
- Label all cases diagnostically and rate relative importance
- Using cardiologists' ratings as outcome standard, would use all available information to develop a model predicting difficulty.
- By having ~600 learners purposefully rate cases, would validate both the CCM and LAA in terms of efficiency and effectiveness.



Accomplishment: At end of year one, 80,000 ECG downloaded from MUSE and EPIC system. All cases "first-pass" labelled with diagnosis and importance ranking. Presentation software is being tested. Completed pilot studies (Confusable and Pairwise Comparisons) currently analyzing data. Start recruitment for cohort study (NYU, UBC, and Mayo Clinic).

Timeline and Cost

Activities CY	Y1-1	Y1-2	Y2-1	Y2-2
Assemble Library				
Develop CCM				
Validate CCM on 500 participants				
Validate LAA on 50 participants				
Estimated Budget ($K)	**$375k**	**$375K**	**$375k**	**$357K**

Updated: 27 October 2017 – Year 1 Report

Goals/Milestones
CY1-1 Goal – Assemble case library
☑ Focus group to verify controlled vocabulary for diagnosis and importance ranking
☑ Assembled corpus of 80,000 cases and associated clinical data
CY1-2 Goal – Develop Case Characteristics Model
☑ Pilot testing on 100 learners (36 NYU + 50 UBC)
☑ Pairwise comparisons (20/20)
☑ Presentation software in beta-testing phase; ECG database in development
CY2-1 Goal – Validate CCM on prospective cohort
☐ Recruit 500 learners from 3 sites to rate 100 cases each (Start Recruitment)
CY2-2 Goal – Develop Learning Adaptation Algorithm
☐ Develop first version of adaptive learning algorithm
☐ Iteratively improve LAA using data from 50 successive learners

Comments/Challenges/Issues/Concerns
Balance of 35.26% of total costs relate to unliquidated obligations for subcontracts, currently we have total of $118,563.65. Pending invoices - UBC ($23,706), KUMC ($65,637), and Mayo Clinic ($29,220.65). Expect to receive full and final invoicing for UBC and Mayo Clinic within 60 days of budget end date which will be 11/29/2017. KUMC requested additional time to fully expend and invoice for Dr. Lineberry's effort. If all unliquidated obligations are deducted from the balance we would expect carryover for only 3.83% with balance of $17,891.93.

Budget Expenditure to Date
Projected Expenditure: $66,129.34 Actual Expenditure: $484,405.96

9. APPENDICES: Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.

APPENDIX 1 - Estimating the difficulty of ECG categories using Rasch Modeling

Pusic, Martin Lineberry, Matthew	NYU School of Medicine Kansas University Medical Center
David Cook Rose Hatala	Mayo Clinic University of B.C.

Background: Some visual diagnoses have greater cognitive complexity than others based, on published models of cognitive task analyses of ECG and other diagnostic imaging. In the table at right, we proposed a straightforward ranking of ECG complexity. A priori we proposed to rank all 120 AHA categories according to this schema so as to a) inform the case characteristics model and b) to provide a basis for selecting example labels for the reaction time determinations. In this study, we sought to validate this ranking through a consensus ranking process.

Level	ECG Feature	Example Diagnosis	Process
1	Segment	Prolonged PR interval (Type 1 Block); Diffuse ST Elevation	Pattern recognition
2	Single Lead	Wolf-Parkinson-White; Brugada	
3	Sextets	Frontal: Ventricular hypertrophy; LBBB; RBBB Axial: Axis deviation	
4	12-Lead	Pericarditis; Myocardial infarction	
5	Two Diagnoses	Myocardial ischemia with LBBB.	Deductive analysis

Procedure: Using an established psychometric process, we had experts order, in terms of their difficulty, successive sets of ECGs. Using pairwise comparisons, they ordered the ECGs in terms of difficulty and specified the variables which make one ECG more difficult than another. These variables, along with the case orderings and a literature review, will allow us to develop a conceptual model of ECG difficulty within and across diagnosis vectors.

The ordinal pairwise ratings process was carried out as follows. Twenty cardiology fellows (NYU Fellows) each ranked 12 ECGs in all possible pairs (66 comparisons per participant) according to their perceived difficulty and documented the variables/terms they used.

Analysis: The resulting dataset was modelled using a Rasch 1-parameter item-response model to generate item locations for each of the 12 ECGs.(See Figure). Scale locations were inferred from the proportions of judgements in favor of that particular case/diagnosis. Groupings by item locations were determined and then interpreted according to the qualitative judgements of the participants. All findings are preliminary.

Results: 20 Cardiology fellows completed all pairwise ratings. The rankings of the diagnoses were considerable different from the *a priori* prediction of cognitive complexity.(See Table 2). In fact, the three cases which were judged to have the lowest cognitive complexity (Level 1 – finding present in all leads and involving on a segment) were in fact judged to be the most difficult diagnoses. On qualitative inspection, there were three natural groupings of diagnoses. The easiest cases involved common diagnoses with widespread and distinct abnormalities (e.g. ST-elevation MI, Atrial Fibrillation) while the most difficult cases involved subtle findings that could be confused with normal (e.g. Prolonged QT syndrome or second degree AV block). Cases with intermediate levels of difficulty suggest findings which require both detection and decoding (e.g. Pericarditis).

Implications: The cognitive complexity of an ECG finding correlated poorly with the actual ranking based on pairwise comparisons by experts. These findings suggest that frequency of diagnosis and magnitude of deviation from normal are relatively more important factors.

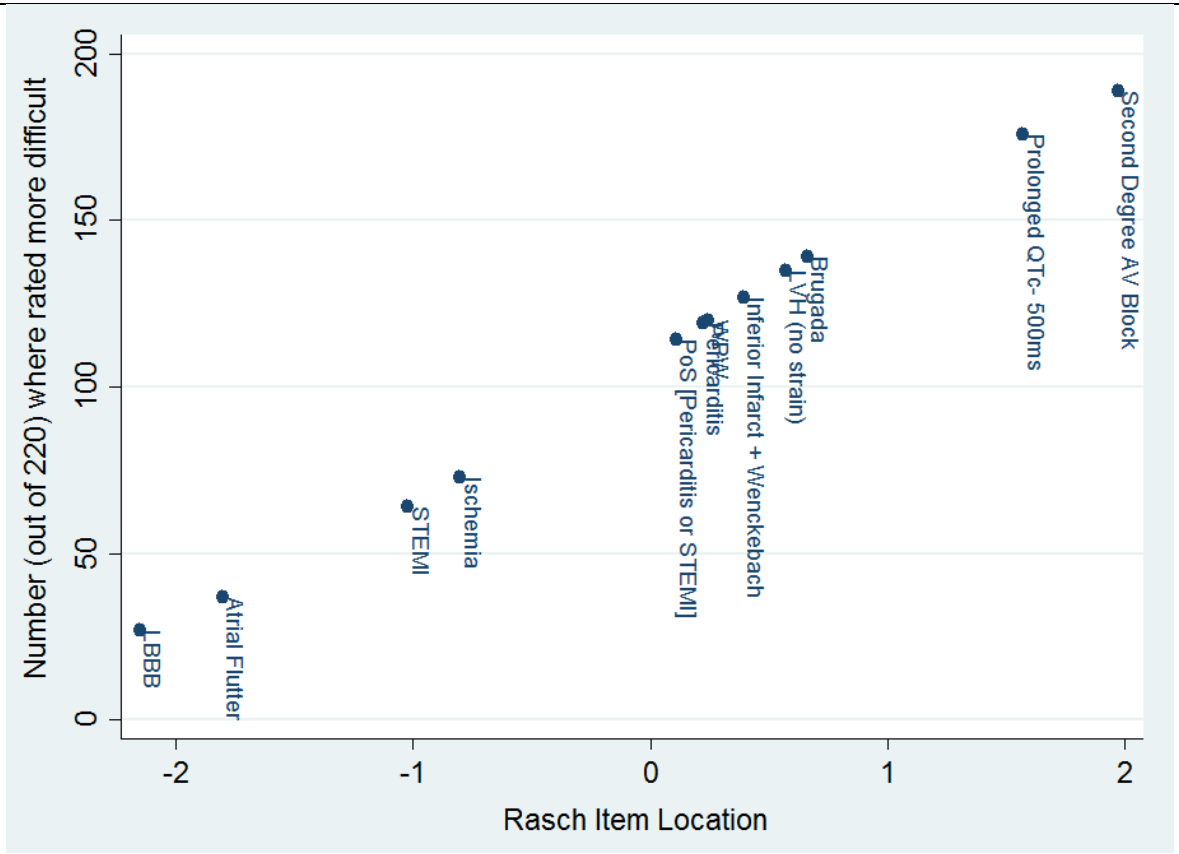
Reference: Baldwin, Peter, Joseph Bernstein, and Howard Wainer. "Hip psychometrics." *Statistics in medicine* 28.17 (2009): 2277-2292.

Table 1 - Diagnoses ranked by pairwise difficulty

Diagnosis	Cognitive Complexity Level	Pairwise - Count Preferred	Rasch Location	Potential Rationale
Second Degree AV Block	1	189	1.974	Subtle finding
Prolonged QTc- 500ms	1	176	1.573	Subtle finding
Brugada	2	139	0.664	Subtle finding
LVH (no strain)	3	135	0.569	Subtle finding
Inferior Infarct + Wenckebach	5	127	0.393	Decoding
WPW	2	120	0.243	Decoding
Pericarditis	4	119	0.222	Decoding
PoS [Pericarditis or STEMI]	5	114	0.109	Decoding
Ischemia	4	73	-0.801	Decoding
STEMI	4	64	-1.021	Common Diagnosis
Atrial Flutter	4	37	-1.802	Common Diagnosis
LBBB	3	27	-2.151	Common Diagnosis

Diagnoses arranged by level of difficulty. For example, the second degree AV block case was judged most difficult as it was ranked more difficult when compared with the other diagnoses 189/220 times. Its Rasch location is a considerable "distance" from the mean of 0. However, the finding itself is not cognitively complex (level 1) as it is a visual recognition task of a single ECG segment.

Figure 2 - Rasch Item Locations



Rasch item locations (along x-axis) show the distances between diagnostic groupings. These grouping support a theory as to why some cases are more difficult than others (see "Potential Rationale" in the Table 1.)

APPENDIX 2 - A Model for Diagnostic Discernment in Visual Diagnosis of ECGs

Pusic, Martin Lineberry, Matthew	NYU School of Medicine Kansas University Medical Center
Julie Friedman Barry Rosenzweig Jeffrey Lorin	NYU School of Medicine

Background: In visual diagnosis, representations of human anatomy or physiology are considered by the clinician in a cognitive process that includes both pattern recognition and analytical deduction in service of accurate therapeutic decisions. In the health professions they can include photographs of physical examination findings (e.g. rashes), radiographs (e.g. for detection of fractures) and electrocardiograms for the detection of cardiac pathology.

Baldwin et al, in a study of orthopedic surgeons classifying radiographs in terms of severity of hip fractures, found that they could indeed develop a model that describes individual orthopedic surgeons' visual diagnostic thresholds by category of severity.(Baldwin) The patient-level importance lays in a surgeon choosing the surgical approach that appropriately matches risk to the level of injury.

Besides severity of illness, another important distinction in clinical decision making lies in the "confusability" of two distinct diagnoses on a particular visual representation. In pilot work, we confirmed that on an electrocardiogram (ECG) performed in the setting of chest pain, clinicians may confuse the diagnosis of pericarditis (inflammation of the membrane surrounding the heart) and ST-elevation myocardial infarction (STEMI; life-threatening blockage of a coronary artery). The former is treated as an outpatient with anti-inflammatory medication while STEMI often requires urgent cardiac catheterization to mitigate a significant risk of death or incapacitation.

In this study, we sought to determine the degree to which a clinician can discriminate between these two confusable diagnoses on the basis of a standard ECG.

Procedure: We selected 20 ECGs (10 pericarditis; 10 STEMI) where some proportion of the respondents had confused pericarditis with STEMI seeking to have a range of "confusability". We had 32 clinicians (8 experts, 24 residents) rank the ECGs on a 5-point scale with the following anchors: "Definitely Pericarditis", "Probably Pericarditis", "Either Pericarditis or STEMI", "Probably STEMI" and "Definitely STEMI". The ECG cases were rated within Qualtrics, an online survey tool. Each participant did each case twice in two blocks of 20 with different order. Participants were not aware the cases were repeated. Participants were NYU School of Medicine Emergency Medicine residents and faculty, representing a broad range of experience and ability in this domain relevant task.

Analysis: The resulting dataset was modelled using Samejima's graded response model to generate individual clinician tracelines showing how they used the categories. Under this model, each individual rater was considered an exchangeable assessment of the degree to which the ECG reflects a prototypical STEMI case, and each ECG as a replication to be located along a continuum of definitive presentation as STEMI by the response pattern across all raters. This reverses the canonical interpretation of stimuli as items and individuals as possessing some level of ability to be assessed. Just as the standard measurement model allows us to determine how an item performs in separating individual examinees by their level of ability, this interpretation allows us to evaluate how each clinician performs in separating ECGs by their degree of resemblance to STEMI. We plot the tracelines for each rater across the 20 ECGs, and estimate the location of each ECG on a continuous latent scale from Definitely Pericarditis to Definitely STEMI.

Results: Qualitative inspection of the participant tracelines revealed interesting developmental and practice variation patterns (Figures). For all participants, we determined the locations corresponding to an estimated 50% probability of endorsement for the most extreme STEMI and the most extreme Pericarditis response used by each participant based on their first attempts on the 20 cases. (For 11 of the 24 junior and senior residents, the most extreme Pericarditis response was considered Probably Pericarditis, as these residents never endorsed the Definitely Pericarditis response.) The locations of these extreme responses were not significantly further separated in the response patterns of the 17 expert physicians and senior residents, with a median separation of 2.08 logits between the extremes, compared to the 15 junior residents, with a median separation of 1.65 logits, (p-value = .74, 95% CI: (-.77, .44). We repeated the analysis on the participants' second pass through the 20 cases with similar results.

Implications: Graded response models may help in assessing the ability of clinicians to discriminate amongst confusable diagnoses, though more investigation is required. Selected cases based on this analysis will be embedded in the prospective data collection on a large number of participants.

Reference: Baldwin, Peter, Joseph Bernstein, and Howard Wainer. "Hip psychometrics." *Statistics in medicine* 28.17 (2009): 2277-2292.

Figure 1 - Item Characteristic Curves by Level of Expertise for ECG Distinction of Pericarditis from ST Elevation MI

The individuals showed differences in their ability and willingness to use the five categories. Preliminary visual inspection of the tracelines suggests raters with more clinical experience (PGY 6 and above) were marginally more likely to use all available response categories and to demonstrate greater separation between the locations of their response category thresholds.

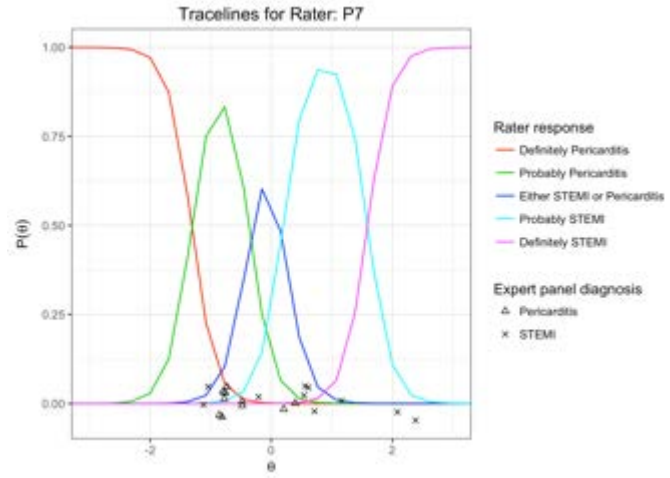
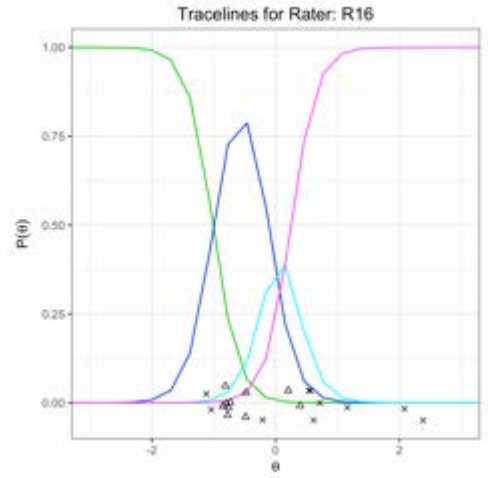
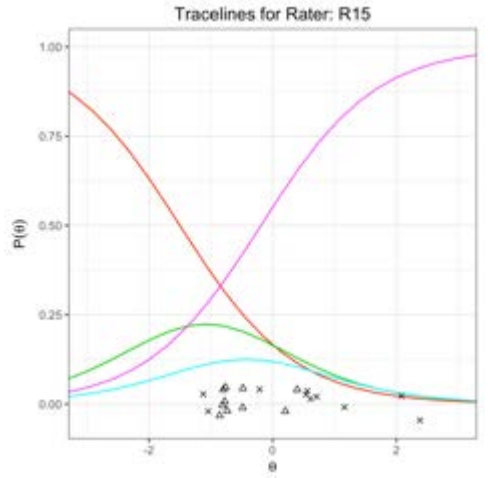
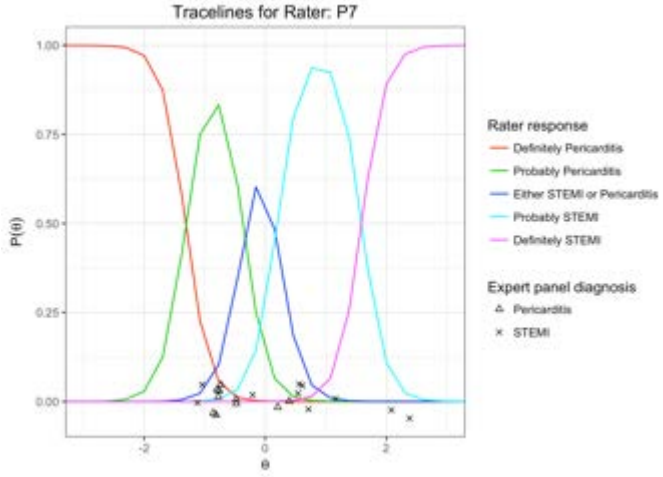
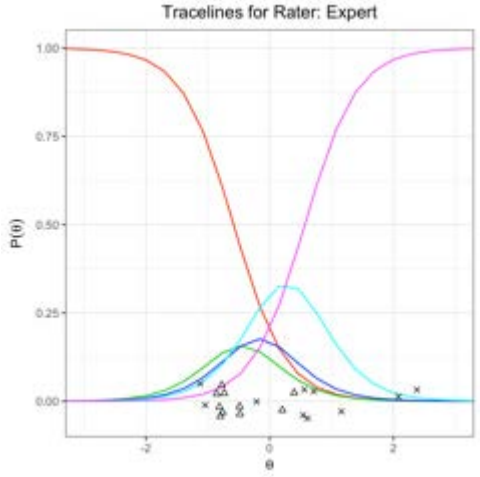
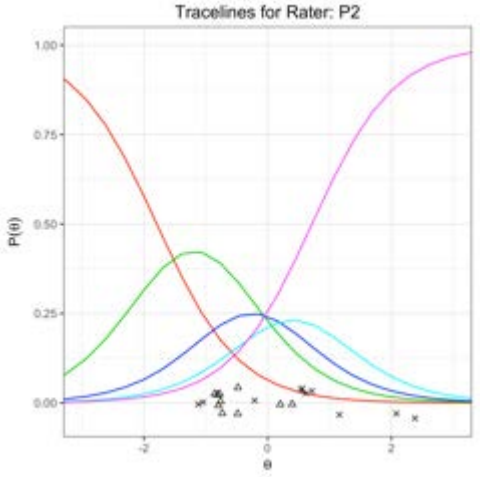
EXPERT	INTERMEDIATE	NOVICE
		
The **expert** uses all 5 of the categories and they each have the intended meaning. Clear separation between Definitely STEMI (pink-rightmost) and Definitely Pericarditis (orange-leftmost)	The **intermediate** uses 4 of the categories but only 3 of them with meaning as the Probably STEMI (blue) is barely distinguished from Definitely STEMI.	The **novice** really has only the two extreme categories and even those overlap considerably.

Figure 2 - Item Characteristic Curves Across Experts for ECG Distinction of Pericarditis from ST Elevation MI

EXPERT	CONSERVATIVE EXPERT -- alters categories --	LIBERAL EXPERT
		
The **balanced expert** uses all 5 of the categories and they each have the intended meaning. Clear separation between Definitely STEMI (pink-rightmost) and Definitely Pericarditis (orange-leftmost)	The **conservative expert** lumps the pericarditis categories together but is careful with the degrees of probability of STEMI.	The **liberal expert** shows the opposite pattern, lumping the STEMI categories but distinguishing 3 levels of pericarditis.

APPENDIX 3 - Expert Ranking of Diagnostic Importance in ECG Interpretation

Pusic, Martin Lineberry, Matthew	NYU School of Medicine Kansas University Medical Center
Sidrah Malik Silas Smith David Cook Rose Hatala	NYU School of Medicine NYU School of Medicine Mayo Clinic University of B.C.

Background: Some ECG diagnoses are important than others. An adaptive tutor, in selecting cases to present to recruits, would ideally take into account case importance. The American Heart Association has developed a controlled taxonomy of diagnostic statements for ECG interpretation. These 120 statements encompass the vast majority of possible diagnoses. In this sub-study, we had experts rank these statements as to their clinical importance for an EM physician.

Procedure: We carried out successive 2-hour focus groups of experts in Emergency Medicine and Cardiology/Internal Medicine including the study cardiologists. The task was to assign relative importance scores to all 120 AHA labels and suggest parameters by which importance might be judged. In a card-sorting exercise, each participant individually chose the 20 “Most Important Diagnoses for an Emergency Medicine clinician” and 20 “Least Important”. Once each clinician had completed this task, they discussed why they made the choices that they did. The focus group was audio-recorded and transcribed for qualitative analysis by a study investigator with experience in these methods.

Analysis: For each diagnostic statement, we calculate the number of times it was listed by a participant as belonging to the 20 most important, 20 least important (deprecated) and intermediate. We report: the numerically most commonly listed 5 “important” diagnoses, the kappa inter-rater agreement between EM physicians and cardiologists, and we list the diagnoses where EM physicians rated the diagnosis amongst the “most-important” and cardiologists listed it in the “deprecated” category, and the converse. Finally, we report commonly occurring qualitative themes.

Results: The five diagnoses listed most commonly across participants were:

- Ventricular tachycardia
- Ventricular fibrillation
- Complete AV Block
- Atrial fibrillation
- Myocardial Infarction

The EM physicians and cardiologists completely agreed on 63% of the diagnoses (expected 33%). Kappa = 0.44. ($p<0.000$). For four diagnoses, the specialists were completely discordant (see Table). Major qualitative themes that emerged included; ECG importance, normal vs. abnormal ECG, chronicle information (prior visit or ECG), clinical context, training level, knowledge gap, unique patterns, impact on patient diagnosis, and ECG educational importance. Some of the sub-themes comprised of; repeated training, enhance technical performance, improve knowledge gap with prompt feedback, and understanding the both low and high impact of ECG interpretation.

		Cardiology		
		Most Important	Intermediate	Deprecated
Emergency Medicine	Most Important	29	10	2 (Hypertrophic Cardiomyopathy, ST Change due to Ventricular Hypertrophy)
	Intermediate	6	25	9
	Deprecated	2 (Sinus Tachycardia, Ectopic Atrial Tachycardia)	16	21

Implications: Emergency medicine physicians and Cardiologists showed good agreement in judging ECG diagnoses as to their importance in the Emergency medicine context. The ratings provide a means for prioritizing adaptive learning of this skill.